

# Leveraging Evolutionary Search to Discover Self-Adaptive and Self-Organizing Cellular Automata

David B. Knoester\*, Heather J. Goldsby†, and Christoph C. Adami\*

\*Dept. of Microbiology and Molecular Genetics

Michigan State University

East Lansing, Michigan 48824

Email: {dk,adami}@msu.edu

†Dept. of Biology

University of Washington

Seattle, Washington 98195

Email: goldsby@uw.edu

**Abstract**—Building self-adaptive and self-organizing (SASO) systems is a challenging problem, in part because SASO principles are not yet well understood and few platforms exist for exploring them. *Cellular automata* (CA) are a well-studied approach to exploring the principles underlying self-organization. A CA comprises a lattice of cells whose states change over time based on a discrete update function. One challenge to developing CA is that the relationship of an update function, which describes the local behavior of each cell, to the global behavior of the entire CA is often unclear. As a result, many researchers have used stochastic search techniques, such as evolutionary algorithms, to automatically discover update functions that produce a desired global behavior. However, these update functions are typically defined in a way that does not provide for self-adaptation. Here we describe an approach to discovering CA update functions that are both self-adaptive and self-organizing. Specifically, we use a novel evolutionary algorithm-based approach to discover finite state machines (FSMs) that implement update functions for CA. We show how this approach is able to evolve FSM-based update functions that perform well on the density classification task for 1-, 2-, and 3-dimensional CA. Moreover, we show that these FSMs are self-adaptive, self-organizing, and highly scalable, often performing well on CA that are orders of magnitude larger than those used to evaluate performance during the evolutionary search. These results demonstrate that CA are a viable platform for studying the integration of self-adaptation and self-organization, and strengthen the case for using evolutionary algorithms as a component of SASO systems.

**Index Terms**—Self-adaptation, self-organization, cellular automata, evolutionary algorithm, finite state machine.

## I. INTRODUCTION

Engineering large-scale distributed computing systems that are resilient in the face of dynamic and often hostile environments is a challenging problem. *Cyber-physical* systems [1], where computing systems interact with the physical world, further exacerbate these difficulties. *Autonomic computing*, so-named for its relation to the autonomic nervous system [2], was envisioned as a means to address these challenges. Frameworks for building such systems are characterized by fault-

tolerance, self-healing, self-organization, and self-adaptivity. Collectively, these are known as the *self-\** properties [3].

While progress has been made on building computational *self-\** systems, nature provides us with many examples of systems that already exhibit some, if not all, of these properties. For example, self-organizing systems can be found in the formation of inorganic crystals and the complex social behaviors of honeybees, ants, and termites [4]. Many organisms also exhibit a form of self-adaptation known as *plasticity*, where their behavior and/or phenotype may change in response to environmental stimuli [5]–[7]. For example, plants may have different flowering times, stem characteristics, or leaf shape based on the resources present in their environment [5]. While many researchers have taken inspiration from biology to develop *self-\** systems, it remains challenging to develop a single computational framework that is *both* self-adaptive and self-organizing (SASO).<sup>1</sup>

*Cellular automata* (CA) are a form of discrete dynamical system where a lattice of  $N$  stateful *cells* change over time. CA were popularized in Conway’s “Game of Life” [9] and rest on a foundation of research established by von Neumann, Burks, and Holland [10], [11], later expanded by Wolfram [12], and have been used to study a variety of phenomena. Cells in a typical CA are always in exactly one of a finite number of states,  $s$ . In most CA (and also here), cell states are binary:  $s = \{0, 1\}$ . The state of all cells in a CA is known as a *configuration*. The initial configuration (time  $t = 0$ ) of a CA is typically random. Configurations change over time based on a discrete *update function*,  $\Phi(n_i)$ , which computes the state

<sup>1</sup>We note that there is some overloading of the term *adaptation*. In biology, *adaptation* refers to how a population changes in a Darwinian manner as a response to selection pressures acting upon it. In evolutionary computation, *self-adaptation* refers to changing parameters (e.g., mutation rate) of an evolutionary algorithm in the midst of a search [8]. In computing systems, *adaptive software* generally refers to software that is able to exhibit a behavioral change at runtime. Here, we use self-adaptation in the computing systems sense.

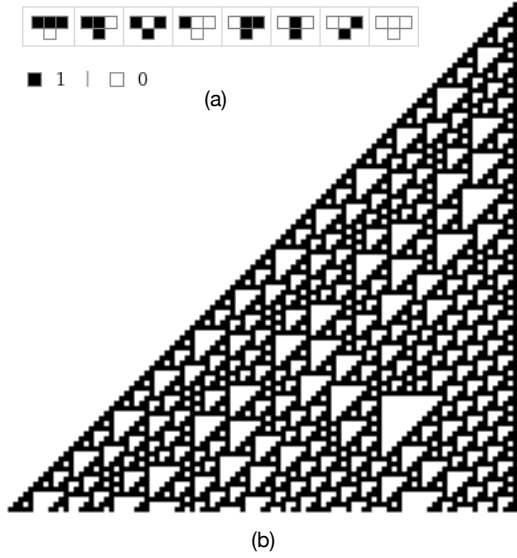


Fig. 1. An elementary CA known as “Rule 110” [12], which has been shown to be Turing complete [13]. (a) The update function for Rule 110. Each possible combination of states for the neighborhood of a cell is given in the top row, while the output (next) state for that cell is given in the bottom row. This update function implements the formula:  $(q \wedge \neg p) \vee (q \oplus r)$ . (b) CA behavior for one initial configuration when updated via Rule 110. Figures from Wolfram|Alpha.

of cell  $c_i$  at time  $t + 1$  based on the states of the cells in its neighborhood  $n_i$  at time  $t$ . Neighborhoods are defined by the topology of the underlying lattice of cells and a *radius*  $r$ , which typically varies from 1 to 3. For example, Figure 1 depicts “Rule 110,” a well-known 1-dimensional (1D) with  $r = 1$ . Here,  $n_i$  consists of the states of three cells:  $\{c_{i-1}, c_i, c_{i+1}\}$  (by convention,  $n_i$  includes the state of cell  $c_i$ ). We note that the boundary conditions of most CA are periodic, i.e., neighborhoods “wrap around” such that a 1D CA is a ring, a 2D CA is a torus, etc.

While update functions for CA often result in intriguing patterns, the relationship of an update function to global behavior is often unclear. Indeed, Land and Belew have proved that perfect solutions to the canonical density classification problem, where a CA must calculate the most common state in its initial configuration, do not exist [14] (this task will be described in more detail in Section III). As a result, many have turned to stochastic search techniques, such as evolutionary algorithms, to generate appropriate update functions for a desired global behavior [15]–[19]. For example, Mitchell et al. have shown that genetic algorithms are capable of discovering update functions for 1D CA with  $N = 149$  [16], while Andre et al. demonstrated a genetic programming approach that discovered high-performing update functions for solving the density classification problem [17]. More recently, Darabos et al. have used evolutionary algorithms to explore the relationship of the update function to the topology of the neighborhood surrounding each cell [15]. While these studies advance our understanding of the principles of self-

organization, they do not include self-adaptive elements.

Here, we describe an approach for studying self-adaptive and self-organizing systems using CA. Specifically, we present a novel evolutionary algorithm that discovers finite state machines (FSMs) that implement CA update functions and allow both self-organization and self-adaptation. We demonstrate these principles by evolving FSM-specified update functions that solve the density classification task in 1-, 2-, and 3-dimensions. We then show that the evolved FSMs are self-organizing and highly scalable, often performing well on CA that are orders of magnitude larger than those used during evolution. Finally, we describe how the evolved FSMs are self-adaptive, in that their behavior is dependent upon internal states that are effectively a simple form of evolved memory. These results support the use of CA as a study system for exploring SASO principles, illustrate one method by which self-adaptation can be integrated into self-organizing systems, and further strengthen the case for the use of evolutionary algorithms as a component of SASO systems.

## II. RELATED WORK

From von Neumann’s early work on self-reproducing automata [10] to Wolfram’s later work on complexity [12], CA have been used to study self-organizing systems across a broad range of domains. For example, CA have been used in image processing [20], to simulate distributed behavior of flocks of birds and schools of fish [21], as model systems for artificial life and parallel processing [22]–[24], and for calculating convex hulls and Voronoi diagrams [25]. Traditionally, CA update functions are specified as a binary lookup table and rely exclusively on the current state of a cell and its neighbors. However, some CA also have the capacity for *memory*, where the update function takes into account the previous state(s) of a cell and the cells in its neighborhood [26]. Other CA replace the typical update function with discrete- or continuous-time artificial neurons, and are known as *cellular neural networks* [27], [28]. Here, to enable self-adaptation, we use an evolutionary algorithm to discover *both* state transition logic and memory characteristics that define a CA update function.

Evolutionary algorithms (EAs) are search techniques inspired by the process of evolution by natural selection [29]. EAs have a long history of being used to solve optimization problems in engineering [30], and as a tool to study and develop self-organizing systems [31]–[35]. A traditional EA comprises a fixed-size population of candidate solutions to a given problem. Each generation, candidate solutions within a population are evaluated to establish how well they solve the problem. Subsequently, a subset of individuals are mutated and recombined to produce new solutions. The next generation of the population is then constructed by selecting individuals from the current generation and newly-produced offspring solutions. Each candidate solution has a genome, which describes its solution to the problem. This genome can be quite simple (a bitstring), or more complex (a tree of Lisp-like *s*-expressions). EAs have previously been applied to the discovery of simple update functions for CA, most

frequently using a genetic algorithm to define a binary rule table [16], [18], [19]. However, there have been approaches that also define more complex update functions and neighborhood topologies [15], [17]. Historically, density classification on 1D CA with  $N = 149$  has been recognized as a demanding problem for genetic algorithms, and a variety of approaches have been evaluated [36]. The work presented here differs in that (a) we present results for 1-, 2-, and 3D CA and (b) the evolutionary algorithm used here is *allowed* to evolve self-adaptation capabilities, but it is by no means certain that it will do so.

*Evolutionary programming* is a kind of EA that focuses on the discovery of finite state machines [37]. Evolutionary programming has frequently been used to address questions in machine intelligence [38]–[41]. In contrast, the approach taken in this paper is inspired by the evolution of *Markov networks*, which are networks of random variables with the Markov property, such that edges between nodes encode arbitrary fuzzy logic gates [42]. Markov networks have been shown to be remarkably effective at exploring questions related to the dynamics of collective behavior, including predator confusion [43] and the formation of the selfish herd [44]. While we do not make use of stochastic behavior in this paper, Fatés has shown that stochastic approaches to the density classification task can perform quite well [45], a subject which we hope to explore in future work.

### III. METHODS

#### A. Density classification task

Density classification is a challenging and yet tractable distributed computation task for exploring emergent behavior within CA [16]. This task requires the CA to sense the density of cells in 0 (or 1) in its initial configuration and modify all cell states accordingly. Specifically, given an initial (random) configuration of cell states  $IC$ , where the possible states  $s = \{0, 1\}$ , the desired global behavior of the CA is to determine whether a threshold density  $\rho_c$  of cells in  $IC$  are 0 or 1. Here (as in [16]) we use  $\rho_c = 1/2$ . For example, if  $\rho_0$  is the density of 1s in a given initial configuration, then the CA should reach a state of all 1s within  $M$  time steps if  $\rho_0 > \rho_c$ , otherwise it should reach a state of all 0s within  $M$  time steps.  $M$  is a parameter that depends on the size of the lattice, and is set to 2 times the number of cells. In this paper, we evolve the update rules for CA that address the density classification task in three different dimensions and topologies (depicted in Figure 2).

#### B. Evolving update functions

Rather than describing the behavior of cells using a binary rule table, here we use an evolutionary algorithm to discover finite state machines that implement update functions for the density classification task. The candidate solutions (FSMs) in this EA are encoded in evolving genomes (described below). As part of the evaluation process, we test each FSM on a set of randomly generated initial conditions ( $IC$ s). Specifically,

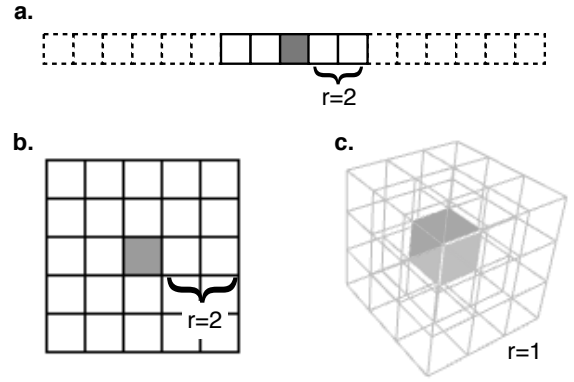


Fig. 2. (a) 1D cellular automata (CA) with radius  $r = 2$ . Each cell in the array determines its state at time  $t + 1$  based on the states of the 5 cells in its neighborhood (itself plus the state of its  $2 * r$  neighbors). (b) 2D CA with radius  $r = 2$ ; each cell determines its state at time  $t + 1$  based on the states of 25 cells. (c) 3D CA with radius  $r = 1$ ; each cell determines its state at time  $t + 1$  based on the states of 27 cells. 2D and 3D topologies both use a Moore neighborhood.

during each evaluation of an  $IC$ , all cells in the CA independently execute their own copy of the genetically encoded FSM. For example, in a 1D CA of 35 cells, there are 35 identical copies of an FSM that has been translated from a genome. During evaluation, the FSMs corresponding to these 35 cells each take input from their neighborhood and produce a single output bit, which is the state of their corresponding cell at time  $t + 1$ .

As in [16], the initial conditions are generated based on a density that is drawn from a uniform distribution  $[0, 1]$ . This density determines the number of 1s present in the  $IC$  (the positions of these 1s are randomly shuffled; all other bits are 0). Finally, we draw  $IC$ s from three different sets: those where the density is strictly less than 0.5, those with density greater than 0.5, and a binomial distribution with  $p = 0.5$ . This latter case is considered the most difficult, as on average there will be an equal number of 0s and 1s in the  $IC$ s.

The *fitness* of a given genome, which is used by the EA to calculate survival and reproduction probabilities, is calculated as follows. First, we translate the genome into an FSM (this process is described in more detail below). This FSM is then evaluated on 100 different initial conditions, and the mean number of correct classifications is calculated. No credit is given for partial solutions. To reduce the chance of an FSM being “lucky” during fitness evaluation, we change the set of initial conditions that are used to evaluate an FSM every generation and average the fitness of a genome over the fitness of its immediate 10 ancestors. This approach is similar to noise mitigation via multiple fitness evaluations [46], but avoids the additional computational overhead.

#### C. Genetic encoding of finite state machines

We base our approach on the evolution of Markov networks [42] and adapt it to evolve FSMs. As shown in Figure 3, we begin by defining a fixed number of binary state variables.

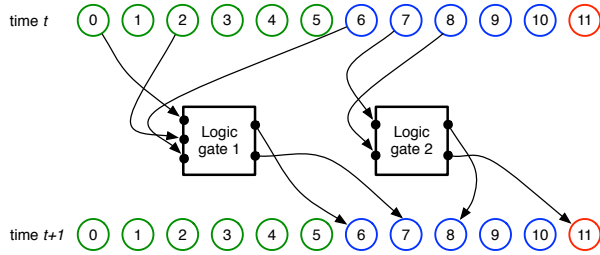


Fig. 3. A simple evolved finite state machine. The inputs and outputs for two logic gates are drawn from the predefined global state variables. Here, gate 1 takes 3 inputs at time  $t$ : 2 from input state variables (green) and 1 from a hidden state variable. Gate 1 also writes into 2 hidden state variables (blue) at time  $t + 1$ . Similarly, gate 2 takes input from 2 hidden states, and writes to both an output state (red) and a hidden state.

Each state variable is either an input that can accept state input from a given cell's neighborhood; an output that produces the next state of a cell; or a hidden state variable that is internal to a cell's FSM and can be used to implement memory. Updates to these state variables are controlled via evolved logic gates that are translated from the genome.

Figure Figure 4 depicts the genome encoding for an FSM. The logic gates that comprise an FSM are embedded in a circular list of integers (the genome) that contains a series of *genes*. Each gene encodes a single logic gate, and defines that gate's fan-in and fan-out (number of inputs and outputs), underlying truth table that defines the logic operation performed by the gate, and the mapping of inputs and outputs to this gate to the global input, output, and hidden state variables. These logic gates operate much like a neuron in an artificial neural network, where all input and output from gates in the FSM is via these global state variables (it may be helpful to think of these state variables as a scoreboard, where each gate synchronously updates a global value). The beginning of a gene is indicated by a *start codon* (a specified sequence of two integers) and the gene's size is determined by the size of the corresponding gate's truth table. The truth table for a given logic gate is defined by the  $2^{n_{in}}$  integers following the output state IDs, where  $n_{in}$  is the number of inputs to that gate. During each time step of an FSM, a given logic gate produces output as follows. First, the inputs to the logic gate are converted in order to a binary number  $r$  that is used to index the corresponding  $r$ 'th integer in the truth table segment of that gene. The  $n_{out}$  low-order bits of that integer are then isolated, where  $n_{out}$  is the number of outputs from that gate, and these binary values are assigned in order to each output state variable.

Table I shows an example evolved logic gate, where  $X$  and  $Y$  are binary inputs,  $Z$  is a binary output. This truth table implements the logic function  $Z = X \vee \neg Y$ , illustrating the ability of evolution to use non-standard logic functions. In all cases, input to— and output from these gates are binary.

Figure 5 is a high-level depiction of an evolved FSM. Here we see three inputs (green ovals; top) that are used by two evolved logic gates (black squares; middle). These inputs are

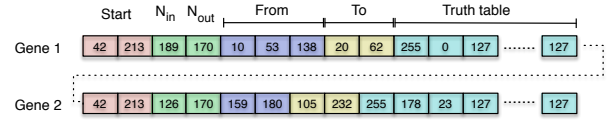


Fig. 4. Example genome with two genes that encode the two gates in Fig. 3. The doublet (42, 213) is the start codon that marks the beginning of a gene (red blocks). The next two codons describe the fan-in and fan-out of the gate (green blocks), and the following codons map the inputs (blue blocks) and outputs (yellow blocks) of that gate to state variables. The remaining codons encode the gate's truth table (cyan blocks). Numeric values in the genome are converted via a mod operation into the appropriate range.

TABLE I  
EXAMPLE TRUTH TABLE FOR AN EVOLVED LOGIC GATE.

X	Y	Z
0	0	1
0	1	0
1	0	1
1	1	1

the states of three neighboring cells. One of these gates (A::L, middle left) produces output that is fed into a hidden state variable (blue oval, bottom left). This hidden state variable is in turn used by the other logic gate (B::L, middle right), and both logic gates produce output for the cell's next state (red oval, bottom right). When multiple gates produce output into the same state variable, the values are logically ORed together.

Each genome in the initial population is a random sequence seeded with 16 randomly generated genes. The entire genome is subject to mutation, including point-mutations (i.e., replacing a codon with a random integer), insertions (i.e., inserting a duplicate of a random sequence from the genome), and deletions (i.e., deleting a random sequence from the genome). Mutations to the underlying genome may alter the truth table, fan-in, and fan-out of the gate, or “rewire” the gate to connect to a different state variables.

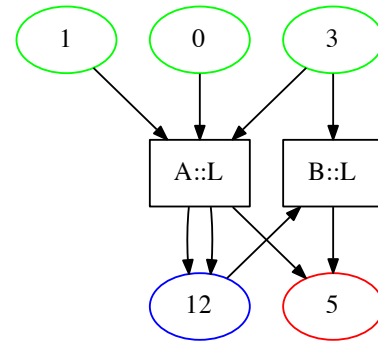


Fig. 5. Example evolved finite state machine. This FSM uses two logic gates (black squares, middle) to accept state input (green ovals, top) from three cells in the CA, and produce the state of the cell at the next time step (red oval, bottom right). One hidden state (blue oval, bottom left) is also used to implement memory. Gates are labeled with a unique code, while state variables are labeled with their index.

TABLE II  
SUMMARY OF EVOLUTIONARY ALGORITHM PARAMETERS.

Parameter	Value
Population size	500
Initial genome size	10,000
Minimum genome size	1,000
Maximum genome size	40,000
Reproduction	asexual
Replacement rate	0.1
Mutation rate (per-site)	0.01
Mutation rate (indel)	0.05
Insertion-deletion size	[16..512]
Number of Initial gates	16
Samples per fitness evaluation	100
Fitness averaging window size	10
Number of updates	10,000
Number of generations (mean)	1000
1D CA topology	$N = 35, r = 2$
2D CA topology	$(M, N) = (7, 7), r = 2$
3D CA topology	$(M, N, P) = (3, 3, 5), r = 1$

#### D. Parameters of the evolutionary algorithm

The specific techniques and parameters that we used to evolve FSMs are summarized in Table II. Each individual in a population of 500 was restricted to a genome of size [1000..40,000] integers, which allows up to a few thousand logic gates (in practice, the largest FSM we have observed here comprises only 120 logic gates). During each update of the EA, the lower 10% of the population was removed via rank-order selection. Parents were selected for asexual reproduction proportional to their fitness until the population reached its initial size, and all offspring undergo mutation. On average, 30 replicates of each treatment were allowed 10,000 updates, which is on average 1,000 generations. Finally, we note that the sizes of the CAs used during evolution are relatively modest; as will be seen, this does not hamper scalability.

### IV. EXPERIMENTS AND RESULTS

#### A. Density classification in 1-, 2-, and 3D

Figure 6 shows the maximum fitness over time on the density classification task for 1-, 2-, and 3D CA (results are averaged across 30 replicates for each treatment; shaded regions indicate 95% confidence intervals constructed via 1,000-fold bootstrapping). Fitness improves rapidly over evolutionary time, eventually reaching a level of approximately 0.95 meaning that, on average, the 95% of the initial configurations are correctly classified (we note that this value is a moving average, as described in Section III-B). This result indicates that our approach to discovering FSM-based update functions for CA using evolvable Markov networks is effective, and can be applied to CA with dramatically different underlying topologies.

To further characterize the performance of CA in classifying initial densities generated by the approach described here, we next analyze the *dominants*, which are the most-fit candidate solutions (i.e., FSM specified update functions) at the end of each replicate (30 dominants per treatment). Figure 7 shows the fraction of correctly classified initial configurations for

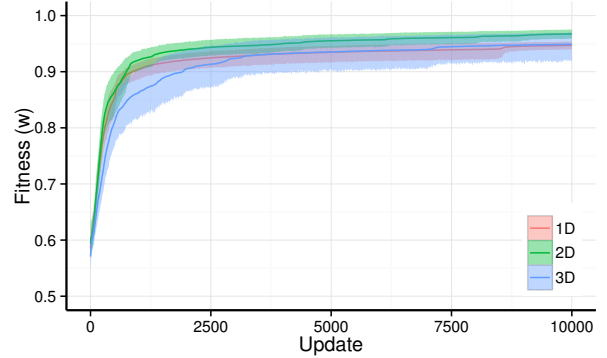


Fig. 6. Average maximum fitness over time for 1-, 2-, and 3D cellular automata solving the density classification problem. Shaded regions are 95% confidence intervals. The 10,000 updates shown here correspond to, on average, 1,000 generations of evolution.

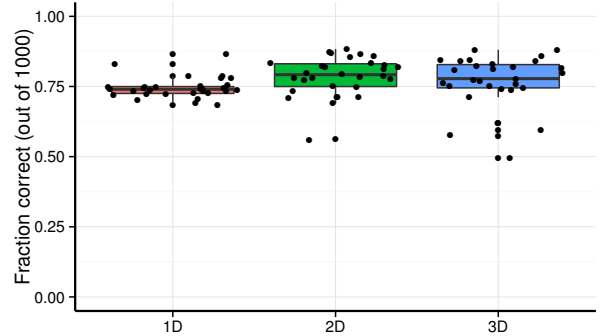


Fig. 7. Boxplot showing the fraction of correctly classified initial configurations (out of 1000) for the dominant evolved solutions from each treatment (30 replicates per treatment). Dots are (jittered) values from each replicate.

each dominant when tested on 1,000 random initial configurations drawn from a binomial distribution ( $p = 0.5$ ). The best dominants from each treatment correctly classified 86.5%, 88.4%, and 88.1% of the initial configurations for the 1-, 2-, and 3D treatments, respectively. This result further demonstrates that evolved update functions perform well even under the challenging 3D topology.

#### B. Example evolved solutions

Figure 8 shows an example of the behavior of a dominant (high fitness) rule drawn from one replicate population of the 1D treatment. We show the state of each cell over time during the evaluation of a single initial configuration. The CA begins with a majority of black cells. For each time step, all cells execute their FSM to determine their next state. In this case, the cells self-organize from an initially random pattern and transition to an all-black configuration. This illustration shows how information from different regions of the configuration are used to produce coherent changes over time.

Similarly, Figure 9 depicts the configurations over time during the evaluation of one 2D initial configuration, and Figure 10 shows configurations for one 3D initial configuration.



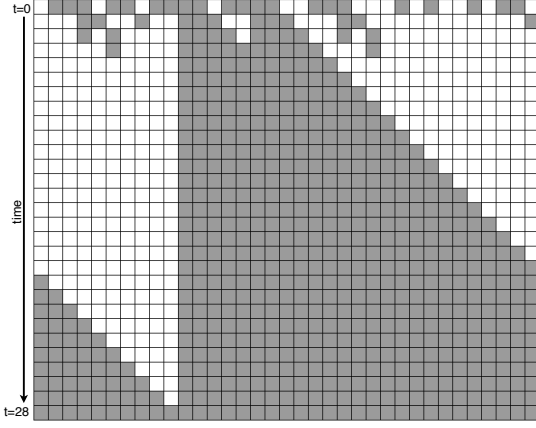


Fig. 8. Example behavior of a 1D CA. Each row in this figure represents the configuration of the CA at a single point in time, with time advancing down (top row is  $t = 0$ ). The final configuration (all cells 1), represents the correct classification for this initial configuration.

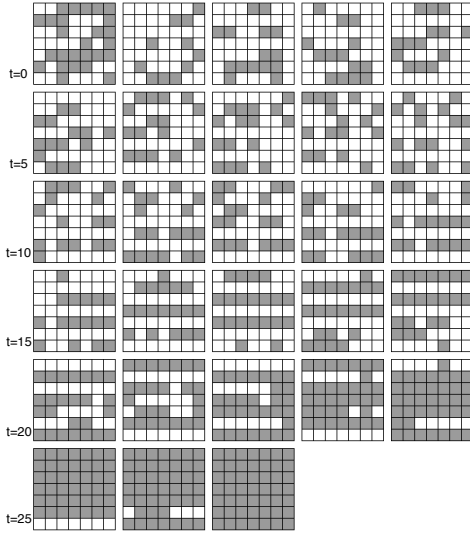


Fig. 9. Example behavior of a 2D CA. Each grid in this figure represents the entire state of the CA at a single point in time, with time advancing left-to-right and top-to-bottom (top left is  $t = 0$ ). The final configuration (all cells 1), represents the correct classification for this initial configuration.

As with the 1D case, both figures show the behavior of the dominant individual drawn from the end of a single replicate population and demonstrate the ability of the cells to self-organize towards the correct decision.

### C. Rule density

An important characteristic of CA update functions for density classification task is that the *rule density*, which is the average frequency of 1s that are output by an update function, should be the same as  $\rho_c$ , the critical density threshold for classification [47]. Thus, we can assess the efficacy of our approach to evolving FSMs to specify update functions by comparing the evolved rule density to the ideal rule density identified by theory. While this is straightforward to calculate for update functions that are based on binary rule tables,

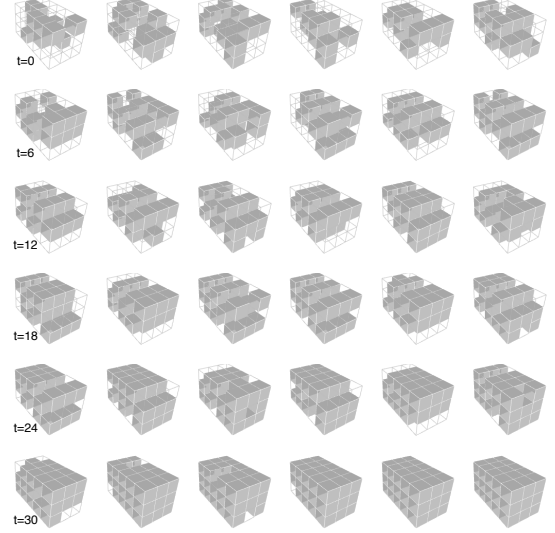


Fig. 10. Example behavior of a 3D CA. Each cuboid in this figure represents the entire state of the CA at a single point in time, with time advancing left-to-right and top-to-bottom (top left is  $t = 0$ ). The final configuration (all cells 1), represents the correct classification for this initial configuration.

calculating the rule density for evolved FSMs requires that we evaluate the FSM for every possible state, and tally the number of 1s that are output. In this study, the FSMs for our 1D CA have 22 binary state variables, yielding  $2^{22}$  possible states (5 binary state variables for the neighborhood inputs, 1 for the output, and 16 available hidden state variables). While we are able to calculate the exact rule density for 1D CA, the 2D and 3D CA have  $2^{42}$  and  $2^{44}$  possible states, respectively. At present, it is computationally infeasible to calculate the exact rule density for 2- and 3D CA, however we were able to estimate rule density by an unbiased sampling of 1 million states.

We show in Figure 11 the fitness of dominant solutions vs. rule density for each of 1-, 2-, and 3D CA. In general, we observe that many of the evolved update functions do indeed have a rule density approaching  $1/2$ , however others do not exhibit this characteristic. Furthermore, we find little correlation between rule density and fitness (Spearman rank correlation,  $\rho = 0.106$ ). We conjecture that the use of memory (made possible by the use of hidden states) alters the optimal density of the rule table, a hypothesis that we hope to explore in future work. Finally, a number of solutions appear to have densities which belie their relatively high fitness. We note that (1) the states over which we calculate rule density are not necessarily the same states that are *actually visited* during evaluation of an initial configuration, and (2) we have observed scenarios where nearly the entire configuration approaches 1, and then in a coordinated fashion turns completely 0 in a single time step. The density for the rule table corresponding to this behavior would be predominately 1, with perhaps a few 0s. This intriguing behavior is only possible in the presence of memory, and also warrants further investigation.

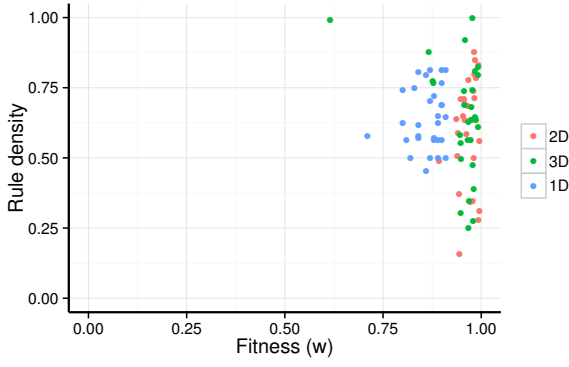


Fig. 11. Rule density density for the evolved dominant solutions from each treatment. 1D results are exact, while 2- and 3D rule densities are estimated from 1M random samples.

#### D. Self-adaptation

Self-adaptation (in the present context) is the unsupervised usage of information that elicits a (run-time) behavioral change in an agent. In general, this requires the capacity to sense environmental characteristics, and to modify internal parameters according to the sensed values. Considering that most CA use an update function defined by a fixed binary rule table (and thus do not include memory), traditional CA are not capable of self-adaptation. However, cells in a CA with memory are in theory capable of using these internal states to change their behavior in response to local stimuli.

The FSMs evolved here include 16 hidden state variables. While we do not predispose our evolutionary algorithm to use any of these hidden states, they could in principle be used as a form of memory for an update function. To determine if memory was an important part of evolved solutions, we evaluated all dominants on 1,000 random (binomial) initial configurations where we held the values of all hidden state variables at 0. We show in Figure 12 how the fraction of correctly classified  $ICs$  ( $\Delta w$ ) changes for each of the 1-, 2-, and 3D CA. As seen there, the average decline in classification accuracy is approximately 0.5 for all treatments. This result strongly indicates that not only are the hidden state variables used, they positively contribute to the performance of evolved FSMs. Because the internal states must depend on the particular  $IC$  the automaton was presented with, we can conclude that the evolved FSMs are self-adaptive.

#### E. Self-organization

Systems where individuals react to each other locally without the availability of global information (as in CA) are examples of *self-organization*. There are various definitions of self-organization, for example: “Self-organization is a process in which pattern at the global level of a system emerges solely from numerous interactions among the lower-level components of the system,” provided by Camazine [48]. While self-organization can be challenging to quantify, we have

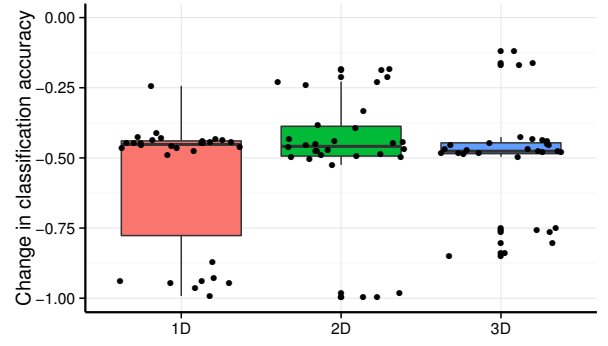


Fig. 12. Boxplot showing the effect of hidden states (memory) on the fraction of correctly classified initial configurations for 1-, 2-, and 3D cellular automata.

previously defined *operational self-organization*,  $S_{op}(x)$ , as:

$$S_{op}(x) = \frac{f(x) - f_{nc}(x)}{f(x) + f_{nc}(x)} \quad (1)$$

where  $f(x)$  is a performance metric of system  $x$  with communication among agents enabled, and  $f_{nc}(x)$  is the same performance metric of  $x$  with communication disabled [33].  $S_{op}(x)$  has the interesting property that values greater than zero indicate the presence of self-organization (communication is beneficial), while values less than or equal to zero indicate the lack of self-organization (communication is harmful or neutral). Here, we can study the degree of self-organization by disabling communication among cells, which is achieved by holding inputs from neighbors (non-self) to 0.

Figure 13 is a boxplot showing  $S_{op}(x)$  for the dominant solutions from each treatment. As shown there,  $S_{op}(x)$  is positive, which indicates that self-organization is always beneficial. Interestingly,  $S_{op}(x)$  is not 1.0; this indicates that in many cases, the evolved FSM correctly classifies an initial configuration without communication among cells. This is an artifact of both our mechanism for disabling communication (holding inputs at 0) and of evolved FSMs having a steady-state behavior that happened to match the correct classification for a given random initial configuration.

#### F. Scalability

An important and desirable characteristic of SASO systems is that they exhibit a high degree of scalability [33]. To explore the scalability of the evolved FSMs, we multiplied each dimension of the 1-, 2-, and 3D CA by a scaling factor  $s = [1..9]$  and re-evaluated the dominant individuals from each treatment on 1,000 new random initial configurations. We can see in Figure 14 the resulting fraction of correctly classified random initial configurations (out of 1,000; we note that the  $x$ -axis is log-scale). We note that the evolved solutions are highly scalable, in some cases scaling multiple orders of magnitude above the input size that they had evolved with, with little decline in performance. Compared to other studies, these results are quite competitive. For example, in [16], Mitchell et

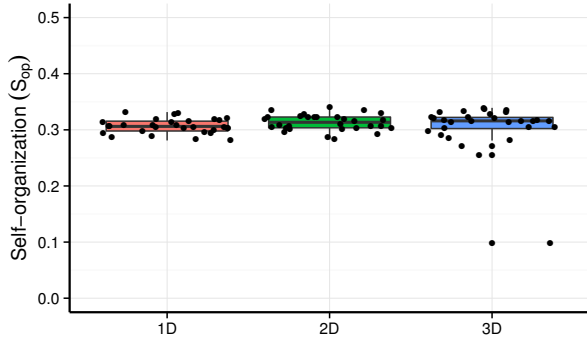


Fig. 13. Boxplot showing operational self-organization for the dominant evolved FSMs from each treatment. Solutions strongly exhibit self-organization.

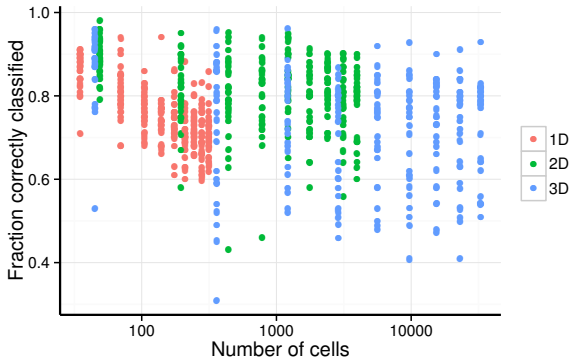


Fig. 14. Each dominant FSM is highly scalable, which we show by testing them in larger CA. We scale each dimension of the 1-, 2-, and 3D CA by a factor  $s$  that ranges from  $[0..9]$ , and plot the fraction of correctly classified configurations vs. the total number of cells. Note that the  $x$ -axis has a log-scale.

al. describe an evolved solution for a 1D  $N = 149$  CA that achieves a classification accuracy of 76.9%, while Andre et al. achieve an accuracy of 82.33%, again in 1D [17]. In contrast, depending on the size and topology of the CA, the solutions we discover here surpass 95%.

Finally, Figure 15 and Figure 16 document the time series of configuration changes for example 1D and 2D CA that have been scaled by 9 and 4, respectively. As seen here, the patterns produced by the evolved FSMs are quite intriguing. While difficult to see within one figure, the animations corresponding to these behaviors indicate that *particles*, that is, cohesive moving blocks of black cells, readily form in these configurations [16].

#### G. Example evolved FSMs

We show in Figure 17 the FSM connection diagrams for each of a 1-, 2-, and 3D CA that achieved fitness scores slightly over 0.99 (the optimal fitness is 1.0). As seen here, these FSMs evolved to use between 3 and 8 hidden and input states. In all cases, the number of inputs used is below that available in each cell's neighborhood, while the memory afforded each cell

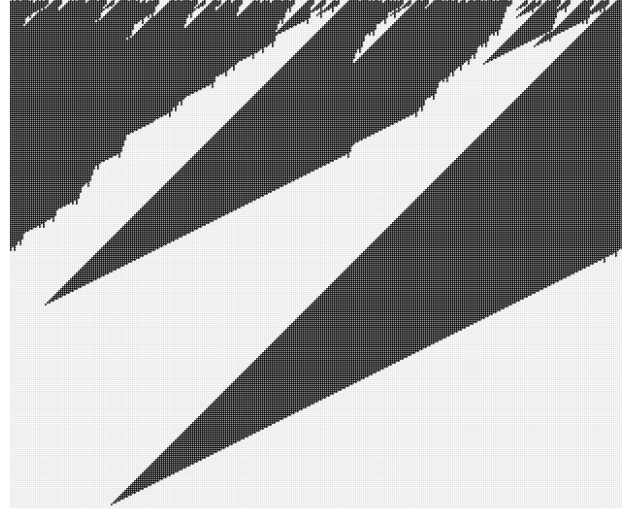


Fig. 15. Example behavior of an evolved FSM on a 1D CA evaluated on a configuration of 315 cells (9 times larger than during evolution).

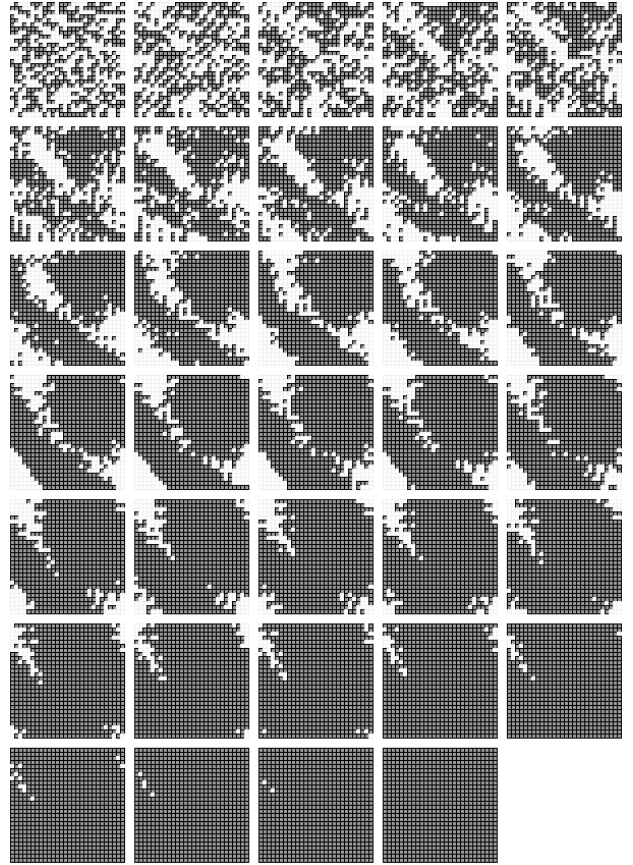


Fig. 16. Example behavior of an evolved FSM on a 2D CA evaluated on a configuration of 784 cells (16 times larger than during evolution).



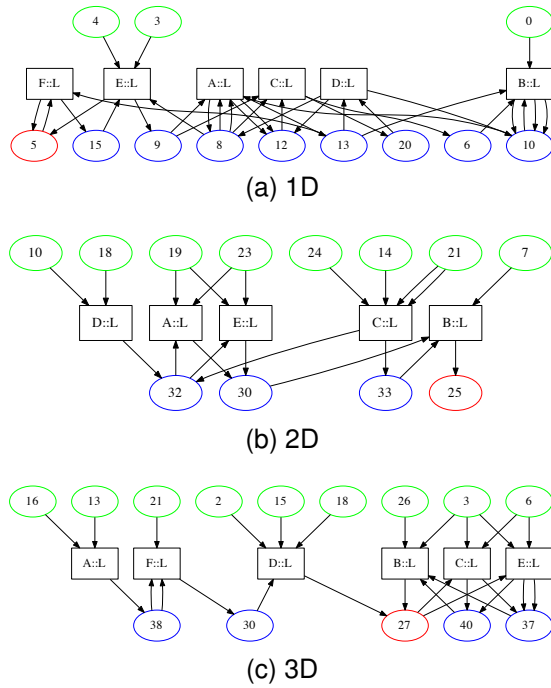


Fig. 17. Evolved FSM connection diagrams for 1-, 2-, and 3D CA. Shown here are the logic gates (black squares), input states (green ovals), and output states (red ovals), and hidden states (blue ovals). Gates are labeled with a unique character for tracking purposes only, while state variables are labeled with their index.

by the inclusion of hidden states improved their classification accuracy. These characteristics of the evolved FSMs highlight the ability of EAs to discover non-intuitive solutions to a challenging problem.

## V. CONCLUSION

We have described an approach in which evolutionary algorithms can be used to discover finite state machines, and used this technique to discover FSMs that perform the density classification task in cellular automata. Moreover, we have shown that CA that use these evolved FSMs are self-organizing, self-adaptive, and highly scalable. In future work, we plan to extend this approach to include other aspects of *self-\** systems, and explore the discovery of FSMs that are fault-tolerant and self-healing.

## REFERENCES

- [1] W. Wolf, "Cyber-physical Systems," *IEEE Computer*, vol. 42, no. 3, 2009.
- [2] J. O. Kephart and D. M. Chess, "The Vision of Autonomic Computing," *IEEE Computer*, vol. 36, pp. 41–50, 2003.
- [3] O. Babaoglu, M. Jelasity, A. Montresor, C. Fetzer, S. Leonardi, A. v. Moorsel, and M. v. Steen, *Self-star Properties in Complex Information Systems: Conceptual and Practical Foundations (Lecture Notes in Computer Science)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.
- [4] B. Hölldobler and E. O. Wilson, *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies*. W. W. Norton & Company, 2008.
- [5] A. D. Bradshaw, "Evolutionary Significance of Phenotypic Plasticity in Plants," *Advances in genetics*, 1965.

- [6] M. J. West-Eberhard, "Phenotypic plasticity and the origins of diversity," *Annual review of ecology and systematics*, 1989.
- [7] A. A. Agrawal, "Phenotypic Plasticity in the Interactions and Evolution of Species," *Science*, vol. 294, no. 5541, pp. 321–326, Oct. 2001.
- [8] O. Kramer, "Evolutionary self-adaptation: a survey of operators and strategy parameters," *Evolutionary Intelligence*, vol. 3, no. 2, pp. 51–65, Feb. 2010.
- [9] M. Gardner, "Mathematical games: The fantastic combinations of John Conway's new solitaire game "life"," *Scientific American*, 1970.
- [10] J. V. Neumann and A. W. Burks, *Theory of Self-Reproducing Automata*. University of Illinois Press, 1966.
- [11] J. H. Holland, "Logical Theory of Adaptive Systems," in *Essays on Cellular Automata*, A. W. Burks, Ed. University of Illinois Press, 1970.
- [12] S. Wolfram, *Cellular Automata and Complexity: Collected Papers*. Addison-Wesley, 1994.
- [13] M. Cook, "Universality in elementary cellular automata," *Complex Systems*, 2004.
- [14] M. Land and R. Belew, "No Perfect Two-State Cellular Automata for Density Classification Exists," *Physical review letters*, vol. 74, no. 25, pp. 5148–5150, Jun. 1995.
- [15] C. Darabos, C. Mackenzie, M. Tomassini, M. Giacobini, and J. Moore, "Cellular Automata Coevolution of Update Functions and Topologies: A Tradeoff between Accuracy and Speed," in *European Conference on Artificial Life 2013*. MIT Press, Sep. 2013, pp. 340–347.
- [16] M. Mitchell, J. P. Crutchfield, and R. Das, "Evolving cellular automata with genetic algorithms: A review of recent work," in *Proceedings of the First International ...*, 1996.
- [17] D. Andre, F. H. Bennett, III, and J. R. Koza, *Discovery by genetic programming of a cellular automata rule that is better than any known rule for the majority classification problem*. MIT Press, Jul. 1996.
- [18] R. Breukelaar and T. Bäck, "Using a genetic algorithm to evolve behavior in multi dimensional cellular automata," in *the 2005 conference*. New York, New York, USA: ACM Press, 2005, p. 107.
- [19] A. Chavoya and Y. Duthen, "Using a genetic algorithm to evolve cellular automata for 2D/3D computational development," ... *of the 8th annual conference on Genetic and ...*, 2006.
- [20] P. L. Rosin, "Training cellular automata for image processing," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2076–2087, 2006.
- [21] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 25–34, Aug. 1987.
- [22] C. G. Langton, "Studying artificial life with cellular automata," *Physica D: Nonlinear Phenomena*, vol. 22, no. 1-3, pp. 120–149, Oct. 1986.
- [23] S. Bandini, G. Mauri, and R. Serra, "Cellular automata: From a theoretical parallel computational model to its application to complex systems," *Parallel Computing*, vol. 27, no. 5, pp. 539–553, Apr. 2001.
- [24] H. Sayama, "Self-Protection and Diversity in Self-Replicating Cellular Automata," *Artificial Life*, vol. 10, no. 1, pp. 83–98, Jan. 2004.
- [25] L. Maignan and F. Gruau, "Convex Hulls on Cellular Spaces: Spatial Computing on Cellular Automata," in *2011 Fifth IEEE Conference on Self-Adaptive and Self-Organizing Systems Workshops (SASOW)*. IEEE, 2011, pp. 67–72.
- [26] R. Alonso-Sanz, "Cellular Automata with Memory," *Computational Complexity: Theory*, 2012.
- [27] L. O. Chua and T. Roska, *Cellular Neural Networks and Visual Computing: Foundations and Applications*. Cambridge University Press, 2002.
- [28] H. Harrer and J. A. Nossek, "Discrete-time cellular neural networks," *International Journal of Circuit Theory and Applications*, vol. 20, no. 5, pp. 453–467, Sep. 1992.
- [29] A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computing*. Springer, 2003.
- [30] D. B. Fogel, "An introduction to simulated evolutionary optimization," *IEEE Transactions on Neural Networks*, vol. 5, no. 1, pp. 3–14, 1994.
- [31] D. B. Knoester, P. K. McKinley, and C. Ofria, "Cooperative Network Construction Using Digital Germplines," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, Jul. 2008, pp. 217–224.

- [32] D. B. Knoester and P. K. McKinley, "Evolution of Probabilistic Consensus in Digital Organisms," in *Proceedings of the IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, Sep. 2009.
- [33] —, "Neuroevolution of Controllers for Self-Organizing Mobile Ad Hoc Networks," in *Proceedings of the IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*. IEEE, Oct. 2011.
- [34] B. E. Beckmann, P. K. McKinley, D. B. Knoester, and C. Ofria, "Evolution of Cooperative Information Gathering in Self-Replicating Digital Organisms," in *Proceedings of the International Conference on Self-Adaptive and Self-Organizing Systems (SASO)*, Jul. 2007.
- [35] H. J. Goldsby, D. B. Knoester, C. Ofria, and B. Kerr, "The Evolutionary Origin of Somatic Cells under the Dirty Work Hypothesis," *PLOS Biology*, vol. 12, no. 5, p. e1001858, May 2014.
- [36] C. Stone and L. Bull, "Evolution of cellular automata with memory: The Density Classification Task," *Biosystems*, vol. 97, no. 2, pp. 108–116, Aug. 2009.
- [37] L. J. Fogel, A. J. Owens, and M. J. Walsh, *Artificial Intelligence through Simulated Evolution*. John Wiley & Sons, Inc., 1966.
- [38] D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. John Wiley & Sons, Inc., 2006.
- [39] G. B. Fogel, "Computational intelligence approaches for pattern discovery in biological systems," *Briefings in Bioinformatics*, vol. 9, no. 4, pp. 307–316, Mar. 2008.
- [40] L. J. Fogel, *Intelligence through simulated evolution: forty years of evolutionary programming*. John Wiley & Sons, Inc., 1999.
- [41] D. B. Fogel, *System Identification through Simulated Evolution: A Machine Learning Approach to Modeling*. Ginn Press, Jul. 1991.
- [42] J. A. Edlund, N. Chaumont, A. Hintze, C. Koch, G. Tononi, and C. Adami, "Integrated Information Increases with Fitness in the Evolution of Animats," *PLoS Computational Biology*, vol. 7, no. 10, p. e1002236, Oct. 2011.
- [43] R. S. Olson, A. Hintze, F. C. Dyer, D. B. Knoester, and C. Adami, "Predator confusion is sufficient to evolve swarming behaviour," *Journal of The Royal Society Interface*, vol. 10, no. 85, Aug. 2013.
- [44] R. S. Olson, D. B. Knoester, and C. Adami, "Critical interplay between density-dependent predation and evolution of the selfish herd," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*. ACM, Jul. 2013.
- [45] N. Fatès, "Stochastic Cellular Automata Solutions to the Density Classification Problem," *Theory of Computing Systems*, vol. 53, no. 2, pp. 223–242, Aug. 2013.
- [46] Y. Jin and J. Branke, "Evolutionary Optimization in Uncertain Environments: A Survey," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 3, pp. 303–317, Jun. 2005.
- [47] M. Capcarrère and M. Sipper, "Necessary conditions for density classification by cellular automata," *Physical Review E*, vol. 64, no. 3, p. 036113, Aug. 2001.
- [48] S. Camazine, *Self-organization in Biological Systems - Google Books*. Princeton University Press, 2003.